Data Preparation for Predictive Modeling

Peggy Brinkmann, FCAS, MAAA
Actuary
Milliman, Inc.

April 1, 2014

Outline

- Why make a big deal about data prep?
- How to do a good data prep
- Case study

What is the big deal?

- You need good prep to meet actuarial Standards of Practice.
- You need good prep to get good results.
- It isn't as easy as you think.

ASOP 23 – Data Quality

3.3 Reliance on Data Supplied by Others

- “the accuracy and comprehensiveness of data supplied by others are the responsibility of those who supply the data”

ASOP 23 – Data Quality

3.5 Review of Data

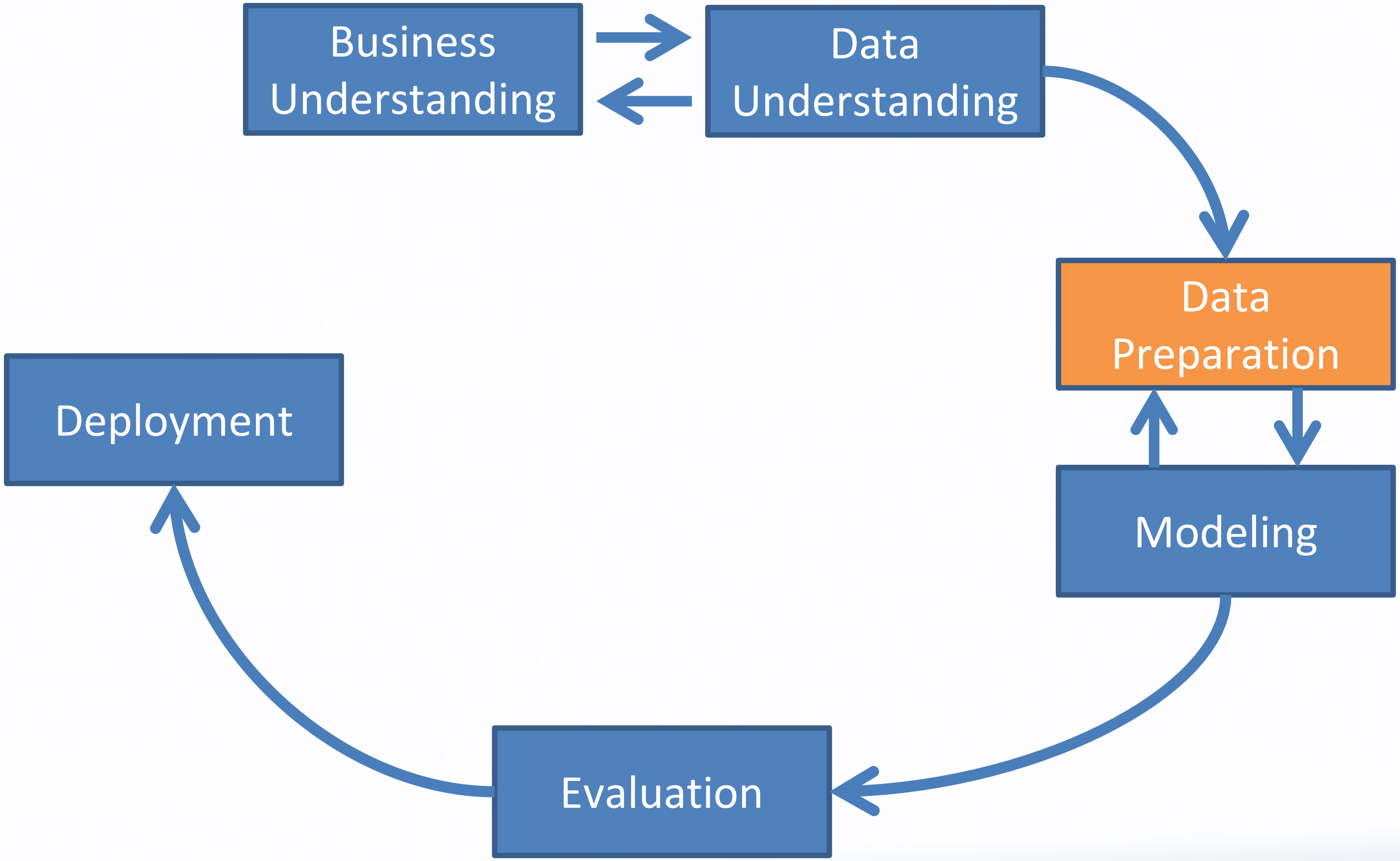
- “the actuary should review the data for reasonableness and consistency, unless, in the actuary’s professional judgment, such review is not necessary or not practical”
- Should consider the following
  - Data definitions
  - Questionable data values
  - Review of prior data

ASOP 23 – Data Quality

3.7 Use of Data

- Is the data sufficient for the analysis?
- Does it require enhancement?
- Are there material defects?
- If the data are inadequate, the actuary should obtain different data or decline the assignment

What can go wrong?



Business Understanding

Data Understanding

Data Preparation

Modeling

Evaluation

Deployment

Steps for good data prep

- Understand the business problem.
- Develop initial analysis plan.
- Review the raw data.
- Calculate the targets and predictors.
- Manually check calculations on examples.
- Review the prepped data.
- Document!

Understanding the business problem

- What is the problem area?
- What is the current solution?
- What are the business goals?
- How will the model(s) be used to achieve the goals?
- What are the implementation constraints?
- What is the desired timeline?

* CASE STUDY *

Problem area	Homeowners insurer in Florida wants to grow its HO3 book of business profitably.
Current solution	Appoint more agencies in profitable areas.
Business goals	Add HO3 policies in Broward, Duval, Lee, Orange, and Volusia counties.
Proposed approach	Build a model to estimate expected loss ratios given data available in a prospect database. Then score the prospect database, and provide list of most profitable prospects to current agencies.
Implementation constraints	Focus on higher-value homes, exclude current policyholders, do not use credit report data.
Timeline	ASAP!

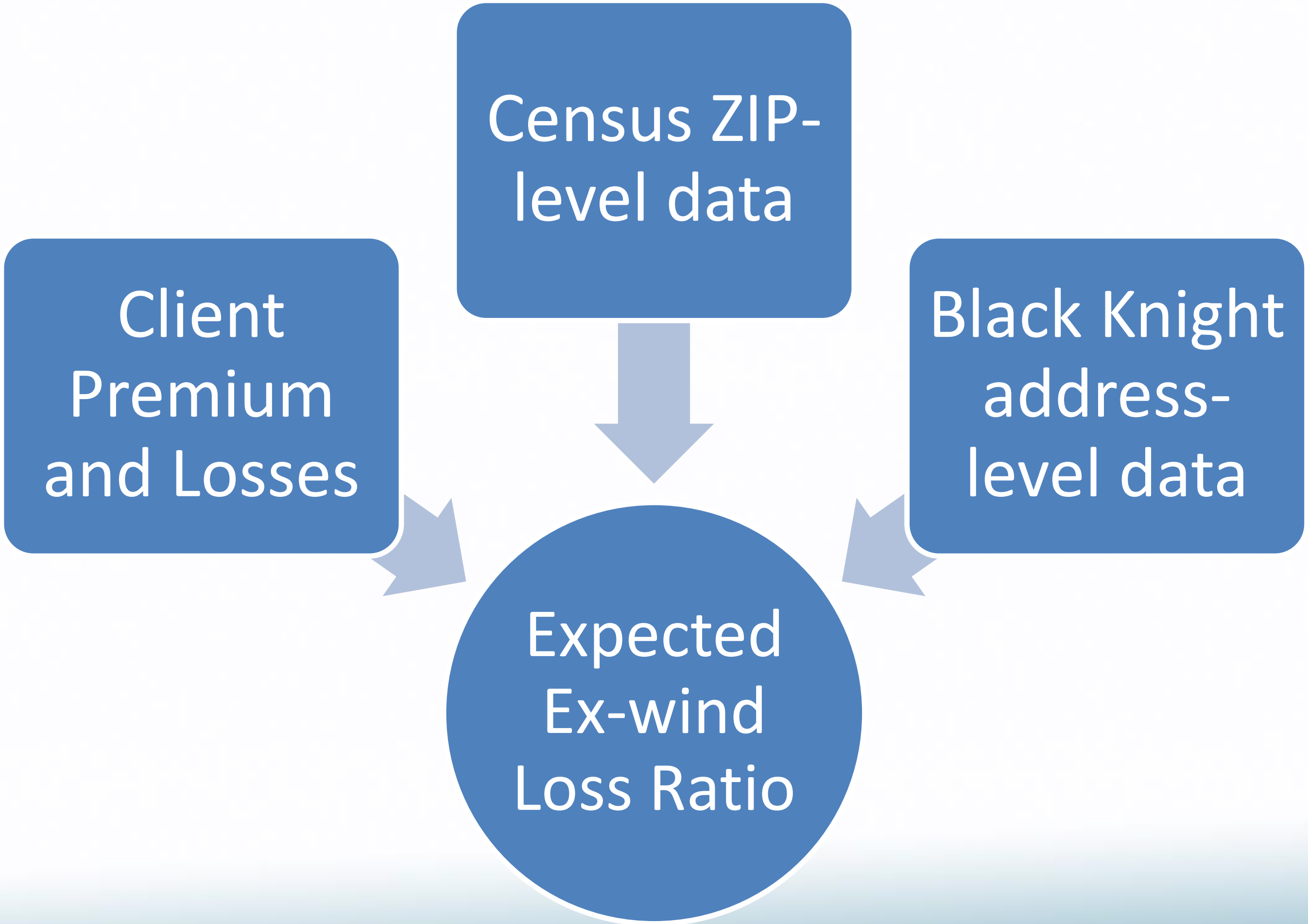
Develop initial analysis plan

- Is the data needed available? Where is it stored, how to access?
- What is the sample size?
- What modeling technique(s) are going to be applied? Using what software?
- How exactly are we defining the target variable(s)? What adjustments will be needed?

* CASE STUDY *

- Need to match prospecting data to client loss ratios. Based on discussions with the client, we used Census data, as well as third party data from Black Knight, who can match onto client data by address
- Client has about 800,000 earned house years
- Use boosted scoring algorithm in EagleEye software
- Target variable will be ex-wind loss ratio at current rates. Data adjustments include
  - Extend exposures to get premium at CRL
  - Exclude wind losses

Data Overview



Census ZIP-level data

Client Premium and Losses

Black Knight address-level data

Expected Ex-wind Loss Ratio

Review the raw data

- Number of records received
- List of fields
- Sample of records
- Distributions of all variables – check for reasonable values
- Produce summary by year

* CASE STUDY *

QC Black Knight.txt

QC Policy Data.txt

QC Claim Data.txt

ASOP 23 questions:

- How is each data element defined?
- Are there questionable data elements?
- What enhancements/adjustments needed?
- Is the data sufficient (i.e. number of potential predictor variables, number of observations)?
- Is it consistent with data used for prior analysis?

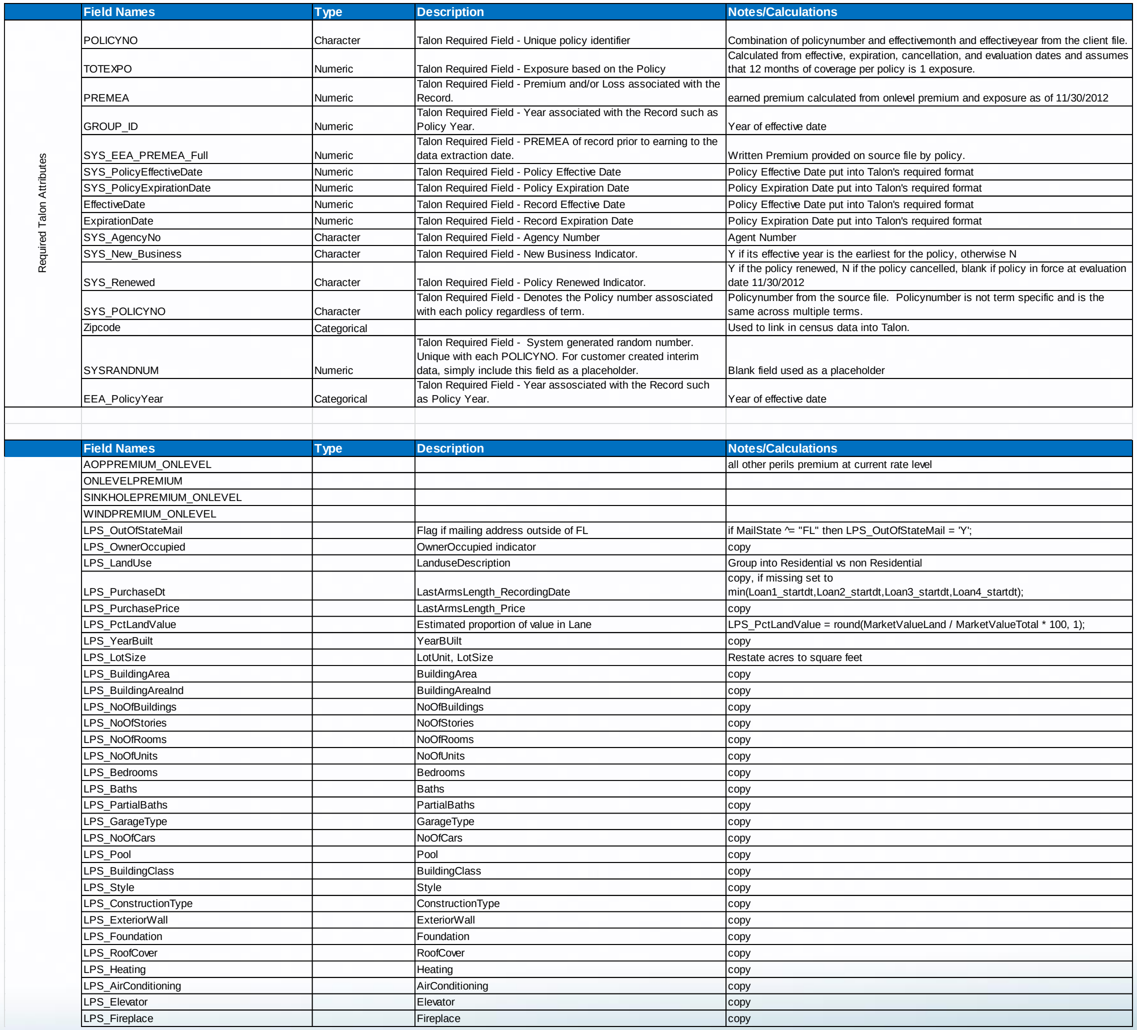
Calculate targets and predictors

- One record per ???
- Definition of target variable
- List and describe of predictor variables
- Handling of missing values, bad data values

* CASE STUDY *

- One record per policy, per policy year
- Target = Ex-wind loss / Ex-wind earned premium
- Missing values – negative 1 for numeric, ~ for categorical

* CASE STUDY *



	Field Names	Type	Description	Notes/Calculations
Required Talon Attributes	POLICYNO	Character	Talon Required Field - Unique policy identifier	Combination of policynumber and effectivemonth and effectiveyear from the client file.
	TOTEXPO	Numeric	Talon Required Field - Exposure based on the Policy	Calculated from effective, expiration, cancellation, and evaluation dates and assumes that 12 months of coverage per policy is 1 exposure.
	PREMEA	Numeric	Talon Required Field - Premium and/or Loss associated with the Record.	earned premium calculated from onlevel premium and exposure as of 11/30/2012
	GROUP_ID	Numeric	Talon Required Field - Year associated with the Record such as Policy Year.	Year of effective date
	SYS_EEA_PREMEA_Full	Numeric	Talon Required Field - PREMEA of record prior to earning to the data extraction date.	Written Premium provided on source file by policy.
	SYS_PolicyEffectiveDate	Numeric	Talon Required Field - Policy Effective Date	Policy Effective Date put into Talon's required format
	SYS_PolicyExpirationDate	Numeric	Talon Required Field - Policy Expiration Date	Policy Expiration Date put into Talon's required format
	EffectiveDate	Numeric	Talon Required Field - Record Effective Date	Policy Effective Date put into Talon's required format
	ExpirationDate	Numeric	Talon Required Field - Record Expiration Date	Policy Expiration Date put into Talon's required format
	SYS_AgencyNo	Character	Talon Required Field - Agency Number	Agent Number
	SYS_New_Business	Character	Talon Required Field - New Business Indicator.	Y if its effective year is the earliest for the policy, otherwise N
	SYS_Renewed	Character	Talon Required Field - Policy Renewed Indicator.	Y if the policy renewed, N if the policy cancelled, blank if policy in force at evaluation date 11/30/2012
	SYS_POLICYNO	Character	Talon Required Field - Denotes the Policy number assosciated with each policy regardless of term.	Policynumber from the source file. Policynumber is not term specific and is the same across multiple terms.
	Zipcode	Categorical		Used to link in census data into Talon.
	SYSRANDNUM	Numeric	Talon Required Field - System generated random number. Unique with each POLICYNO. For customer created interim data, simply include this field as a placeholder.	Blank field used as a placeholder
	EEA_PolicyYear	Categorical	Talon Required Field - Year assosciated with the Record such as Policy Year.	Year of effective date

	Field Names	Type	Description	Notes/Calculations
	AOPPREMIUM_ONLEVEL			all other perils premium at current rate level
	ONLEVELPREMIUM			
	SINKHOLEPREMIUM_ONLEVEL			
	WINDPREMIUM_ONLEVEL			
	LPS_OutOfStateMail		Flag if mailing address outside of FL	if MailState ^= "FL" then LPS_OutOfStateMail = 'Y';
	LPS_OwnerOccupied		OwnerOccupied indicator	copy
	LPS_LandUse		LanduseDescription	Group into Residential vs non Residential
	LPS_PurchaseDt		LastArmsLength_RecordingDate	copy, if missing set to min(Loan1_startdt,Loan2_startdt,Loan3_startdt,Loan4_startdt);
	LPS_PurchasePrice		LastArmsLength_Price	copy
	LPS_PctLandValue		Estimated proportion of value in Lane	LPS_PctLandValue = round(MarketValueLand / MarketValueTotal * 100, 1);
	LPS_YearBuilt		YearBUilt	copy
	LPS_LotSize		LotUnit, LotSize	Restate acres to square feet
	LPS_BuildingArea		BuildingArea	copy
	LPS_BuildingAreaInd		BuildingAreaInd	copy
	LPS_NoOfBuildings		NoOfBuildings	copy
	LPS_NoOfStories		NoOfStories	copy
	LPS_NoOfRooms		NoOfRooms	copy
	LPS_NoOfUnits		NoOfUnits	copy
	LPS_Bedrooms		Bedrooms	copy
	LPS_Baths		Baths	copy
	LPS_PartialBaths		PartialBaths	copy
	LPS_GarageType		GarageType	copy
	LPS_NoOfCars		NoOfCars	copy
	LPS_Pool		Pool	copy
	LPS_BuildingClass		BuildingClass	copy
	LPS_Style		Style	copy
	LPS_ConstructionType		ConstructionType	copy
	LPS_ExteriorWall		ExteriorWall	copy
	LPS_Foundation		Foundation	copy
	LPS_RoofCover		RoofCover	copy
	LPS_Heating		Heating	copy
	LPS_AirConditioning		AirConditioning	copy
	LPS_Elevator		Elevator	copy
	LPS_Fireplace		Fireplace	copy

Manual checks

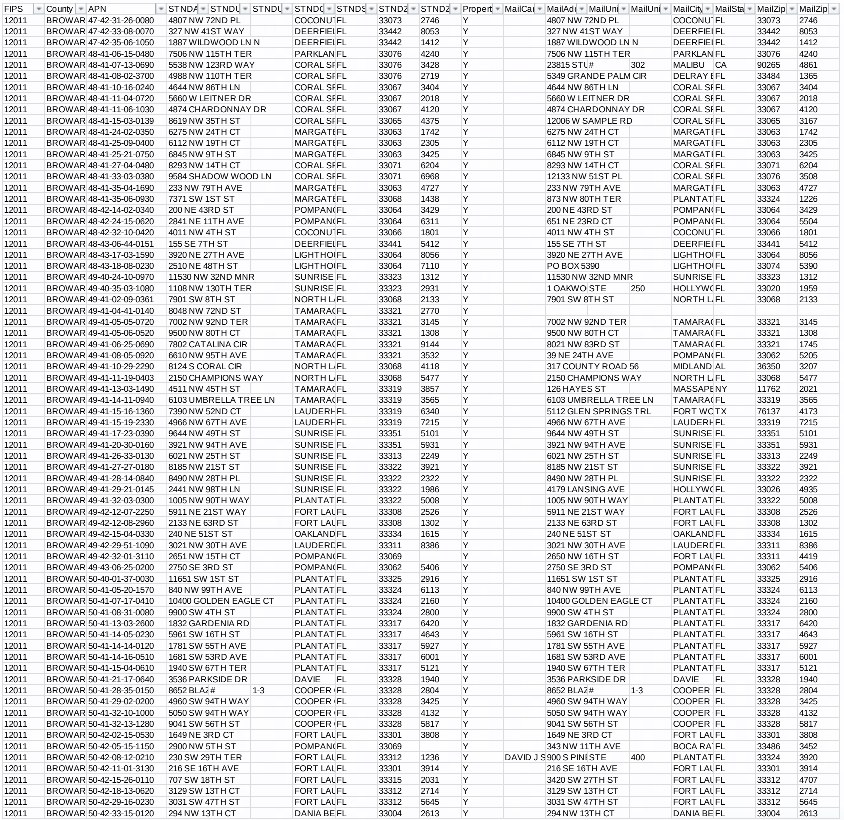
- Select sample for checking – same policies from raw and prepped data
- Manually calculate target variables (e.g. earned exposures, earned premium, incurred losses, renewal) from raw data and compare to prepped
- Manually verify/calculate predictor variables from raw data

* CASE STUDY *

- Checks for onlevel premium calculation
- Checks for exposure and loss fields
- Checks for predictor variable fields

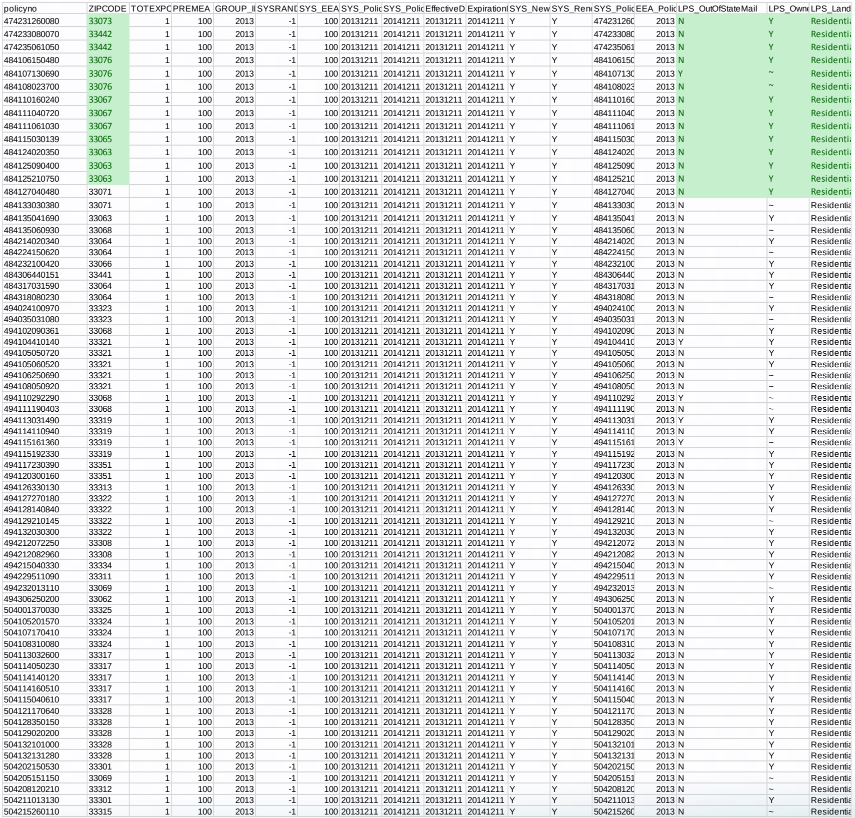
* CASE STUDY *

Raw.xlsx



FIPS	County	APN	STNDA	STNDU	STNDU	STNDC	STNDS	STNDZ	STNDZ	Propert	MailCar	MailAd	MailUni	MailUni	MailCity	MailSta	MailZip	MailZip
12011	BROWAR	47-42-31-26-0080	4807 NW 72ND PL			COCONUT	FL	33073	2746	Y		4807 NW 72ND PL			COCONUT	FL	33073	2746
12011	BROWAR	47-42-33-08-0070	327 NW 41ST WAY			DEERFIEL	FL	33442	8053	Y		327 NW 41ST WAY			DEERFIEL	FL	33442	8053
12011	BROWAR	47-42-35-06-1050	1887 WILDWOOD LN N			DEERFIEL	FL	33442	1412	Y		1887 WILDWOOD LN N			DEERFIEL	FL	33442	1412
12011	BROWAR	48-41-06-15-0480	7506 NW 115TH TER			PARKLAN	FL	33076	4240	Y		7506 NW 115TH TER			PARKLAN	FL	33076	4240
12011	BROWAR	48-41-07-13-0690	5538 NW 123RD WAY			CORAL SP	FL	33076	3428	Y		23815 STU	#	302	MALIBU	CA	90265	4861
12011	BROWAR	48-41-08-02-3700	4988 NW 110TH TER			CORAL SP	FL	33076	2719	Y		5349 GRANDE PALM CIR			DELRAY B	FL	33484	1365
12011	BROWAR	48-41-10-16-0240	4644 NW 86TH LN			CORAL SP	FL	33067	3404	Y		4644 NW 86TH LN			CORAL SP	FL	33067	3404
12011	BROWAR	48-41-11-04-0720	5660 W LEITNER DR			CORAL SP	FL	33067	2018	Y		5660 W LEITNER DR			CORAL SP	FL	33067	2018
12011	BROWAR	48-41-11-06-1030	4874 CHARDONNAY DR			CORAL SP	FL	33067	4120	Y		4874 CHARDONNAY DR			CORAL SP	FL	33067	4120
12011	BROWAR	48-41-15-03-0139	8619 NW 35TH ST			CORAL SP	FL	33065	4375	Y		12006 W SAMPLE RD			CORAL SP	FL	33065	3167
12011	BROWAR	48-41-24-02-0350	6275 NW 24TH CT			MARGATE	FL	33063	1742	Y		6275 NW 24TH CT			MARGATE	FL	33063	1742
12011	BROWAR	48-41-25-09-0400	6112 NW 19TH CT			MARGATE	FL	33063	2305	Y		6112 NW 19TH CT			MARGATE	FL	33063	2305
12011	BROWAR	48-41-25-21-0750	6845 NW 9TH ST			MARGATE	FL	33063	3425	Y		6845 NW 9TH ST			MARGATE	FL	33063	3425
12011	BROWAR	48-41-27-04-0480	8293 NW 14TH CT			CORAL SP	FL	33071	6204	Y		8293 NW 14TH CT			CORAL SP	FL	33071	6204
12011	BROWAR	48-41-33-03-0380	9584 SHADOW WOOD LN			CORAL SP	FL	33071	6968	Y		12133 NW 51ST PL			CORAL SP	FL	33076	3508
12011	BROWAR	48-41-35-04-1690	233 NW 79TH AVE			MARGATE	FL	33063	4727	Y		233 NW 79TH AVE			MARGATE	FL	33063	4727
12011	BROWAR	48-41-35-06-0930	7371 SW 1ST ST			MARGATE	FL	33068	1438	Y		873 NW 80TH TER			PLANTAT	FL	33324	1226
12011	BROWAR	48-42-14-02-0340	200 NE 43RD ST			POMPANO	FL	33064	3429	Y		200 NE 43RD ST			POMPANO	FL	33064	3429
12011	BROWAR	48-42-24-15-0620	2841 NE 11TH AVE			POMPANO	FL	33064	6311	Y		651 NE 23RD CT			POMPANO	FL	33064	5504
12011	BROWAR	48-42-32-10-0420	4011 NW 4TH ST			COCONUT	FL	33066	1801	Y		4011 NW 4TH ST			COCONUT	FL	33066	1801
12011	BROWAR	48-43-06-44-0151	155 SE 7TH ST			DEERFIEL	FL	33441	5412	Y		155 SE 7TH ST			DEERFIEL	FL	33441	5412
12011	BROWAR	48-43-17-03-1590	3920 NE 27TH AVE			LIGHTHOU	FL	33064	8056	Y		3920 NE 27TH AVE			LIGHTHOU	FL	33064	8056
12011	BROWAR	48-43-18-08-0230	2510 NE 48TH ST			LIGHTHOU	FL	33064	7110	Y		PO BOX 5390			LIGHTHOU	FL	33074	5390
12011	BROWAR	49-40-24-10-0970	11530 NW 32ND MNR			SUNRISE	FL	33323	1312	Y		11530 NW 32ND MNR			SUNRISE	FL	33323	1312
12011	BROWAR	49-40-35-03-1080	1108 NW 130TH TER			SUNRISE	FL	33323	2931	Y		1 OAKWO	STE	250	HOLLYWO	FL	33020	1959
12011	BROWAR	49-41-02-09-0361	7901 SW 8TH ST			NORTH LA	FL	33068	2133	Y		7901 SW 8TH ST			NORTH LA	FL	33068	2133
12011	BROWAR	49-41-04-41-0140	8048 NW 72ND ST			TAMARAC	FL	33321	2770	Y								
12011	BROWAR	49-41-05-05-0720	7002 NW 92ND TER			TAMARAC	FL	33321	3145	Y		7002 NW 92ND TER			TAMARAC	FL	33321	3145
12011	BROWAR	49-41-05-06-0520	9500 NW 80TH CT			TAMARAC	FL	33321	1308	Y		9500 NW 80TH CT			TAMARAC	FL	33321	1308
12011	BROWAR	49-41-06-25-0690	7802 CATALINA CIR			TAMARAC	FL	33321	9144	Y		8021 NW 83RD ST			TAMARAC	FL	33321	1745
12011	BROWAR	49-41-08-05-0920	6610 NW 95TH AVE			TAMARAC	FL	33321	3532	Y		39 NE 24TH AVE			POMPANO	FL	33062	5205
12011	BROWAR	49-41-10-29-2290	8124 S CORAL CIR			NORTH LA	FL	33068	4118	Y		317 COUNTY ROAD 56			MIDLAND	AL	36350	3207
12011	BROWAR	49-41-11-19-0403	2150 CHAMPIONS WAY			NORTH LA	FL	33068	5477	Y		2150 CHAMPIONS WAY			NORTH LA	FL	33068	5477
12011	BROWAR	49-41-13-03-1490	4511 NW 45TH ST			TAMARAC	FL	33319	3857	Y		126 HAYES ST			MASSAPE	NY	11762	2021
12011	BROWAR	49-41-14-11-0940	6103 UMBRELLA TREE LN			TAMARAC	FL	33319	3565	Y		6103 UMBRELLA TREE LN			TAMARAC	FL	33319	3565
12011	BROWAR	49-41-15-16-1360	7390 NW 52ND CT			LAUDERH	FL	33319	6340	Y		5112 GLEN SPRINGS TRL			FORT WO	TX	76137	4173
12011	BROWAR	49-41-15-19-2330	4966 NW 67TH AVE			LAUDERH	FL	33319	7215	Y		4966 NW 67TH AVE			LAUDERH	FL	33319	7215
12011	BROWAR	49-41-17-23-0390	9644 NW 49TH ST			SUNRISE	FL	33351	5101	Y		9644 NW 49TH ST			SUNRISE	FL	33351	5101
12011	BROWAR	49-41-20-30-0160	3921 NW 94TH AVE			SUNRISE	FL	33351	5931	Y		3921 NW 94TH AVE			SUNRISE	FL	33351	5931
12011	BROWAR	49-41-26-33-0130	6021 NW 25TH ST			SUNRISE	FL	33313	2249	Y		6021 NW 25TH ST			SUNRISE	FL	33313	2249
12011	BROWAR	49-41-27-27-0180	8185 NW 21ST ST			SUNRISE	FL	33322	3921	Y		8185 NW 21ST ST			SUNRISE	FL	33322	3921
12011	BROWAR	49-41-28-14-0840	8490 NW 28TH PL			SUNRISE	FL	33322	2322	Y		8490 NW 28TH PL			SUNRISE	FL	33322	2322
12011	BROWAR	49-41-29-21-0145	2441 NW 98TH LN			SUNRISE	FL	33322	1986	Y		4179 LANSING AVE			HOLLYWO	FL	33026	4935
12011	BROWAR	49-41-32-03-0300	1005 NW 90TH WAY			PLANTAT	FL	33322	5008	Y		1005 NW 90TH WAY			PLANTAT	FL	33322	5008
12011	BROWAR	49-42-12-07-2250	5911 NE 21ST WAY			FORT LAU	FL	33308	2526	Y		5911 NE 21ST WAY			FORT LAU	FL	33308	2526
12011	BROWAR	49-42-12-08-2960	2133 NE 63RD ST			FORT LAU	FL	33308	1302	Y		2133 NE 63RD ST			FORT LAU	FL	33308	1302
12011	BROWAR	49-42-15-04-0330	240 NE 51ST ST			OAKLAND	FL	33334	1615	Y		240 NE 51ST ST			OAKLAND	FL	33334	1615
12011	BROWAR	49-42-29-51-1090	3021 NW 30TH AVE			LAUDERD	FL	33311	8386	Y		3021 NW 30TH AVE			LAUDERD	FL	33311	8386
12011	BROWAR	49-42-32-01-3110	2651 NW 15TH CT			POMPANO	FL	33069		Y		2650 NW 16TH ST			FORT LAU	FL	33311	4419
12011	BROWAR	49-43-06-25-0200	2750 SE 3RD ST			POMPANO	FL	33062	5406	Y		2750 SE 3RD ST			POMPANO	FL	33062	5406
12011	BROWAR	50-40-01-37-0030	11651 SW 1ST ST			PLANTAT	FL	33325	2916	Y		11651 SW 1ST ST			PLANTAT	FL	33325	2916
12011	BROWAR	50-41-05-20-1570	840 NW 99TH AVE			PLANTAT	FL	33324	6113	Y		840 NW 99TH AVE			PLANTAT	FL	33324	6113
12011	BROWAR	50-41-07-17-0410	10400 GOLDEN EAGLE CT			PLANTAT	FL	33324	2160	Y		10400 GOLDEN EAGLE CT			PLANTAT	FL	33324	2160
12011	BROWAR	50-41-08-31-0080	9900 SW 4TH ST			PLANTAT	FL	33324	2800	Y		9900 SW 4TH ST			PLANTAT	FL	33324	2800
12011	BROWAR	50-41-13-03-2600	1832 GARDENIA RD			PLANTAT	FL	33317	6420	Y		1832 GARDENIA RD			PLANTAT	FL	33317	6420
12011	BROWAR	50-41-14-05-0230	5961 SW 16TH ST			PLANTAT	FL	33317	4643	Y		5961 SW 16TH ST			PLANTAT	FL	33317	4643
12011	BROWAR	50-41-14-14-0120	1781 SW 55TH AVE			PLANTAT	FL	33317	5927	Y		1781 SW 55TH AVE			PLANTAT	FL	33317	5927
12011	BROWAR	50-41-14-16-0510	1681 SW 53RD AVE			PLANTAT	FL	33317	6001	Y		1681 SW 53RD AVE			PLANTAT	FL	33317	6001
12011	BROWAR	50-41-15-04-0610	1940 SW 67TH TER			PLANTAT	FL	33317	5121	Y		1940 SW 67TH TER			PLANTAT	FL	33317	5121
12011	BROWAR	50-41-21-17-0640	3536 PARKSIDE DR			DAVIE	FL	33328	1940	Y		3536 PARKSIDE DR			DAVIE	FL	33328	1940
12011	BROWAR	50-41-28-35-0150	8652 BLAZ	#	1-3	COOPER C	FL	33328	2804	Y		8652 BLAZ	#	1-3	COOPER C	FL	33328	2804
12011	BROWAR	50-41-29-02-0200	4960 SW 94TH WAY			COOPER C	FL	33328	3425	Y		4960 SW 94TH WAY			COOPER C	FL	33328	3425
12011	BROWAR	50-41-32-10-1000	5050 SW 94TH WAY			COOPER C	FL	33328	4132	Y		5050 SW 94TH WAY			COOPER C	FL	33328	4132
12011	BROWAR	50-41-32-13-1280	9041 SW 56TH ST			COOPER C	FL	33328	5817	Y		9041 SW 56TH ST			COOPER C	FL	33328	5817
12011	BROWAR	50-42-02-15-0530	1649 NE 3RD CT			FORT LAU	FL	33301	3808	Y		1649 NE 3RD CT			FORT LAU	FL	33301	3808
12011	BROWAR	50-42-05-15-1150	2900 NW 5TH ST			POMPANO	FL	33069		Y		343 NW 11TH AVE			BOCA RAT	FL	33486	3452
12011	BROWAR	50-42-08-12-0210	230 SW 29TH TER			FORT LAU	FL	33312	1236	Y	DAVID J S	900 S PINE	STE	400	PLANTAT	FL	33324	3920
12011	BROWAR	50-42-11-01-3130	216 SE 16TH AVE			FORT LAU	FL	33301	3914	Y		216 SE 16TH AVE			FORT LAU	FL	33301	3914
12011	BROWAR	50-42-15-26-0110	707 SW 18TH ST			FORT LAU	FL	33315	2031	Y		3420 SW 27TH ST			FORT LAU	FL	33312	4707
12011	BROWAR	50-42-18-13-0620	3129 SW 13TH CT			FORT LAU	FL	33312	2714	Y		3129 SW 13TH CT			FORT LAU	FL	33312	2714
12011	BROWAR	50-42-29-16-0230	3031 SW 47TH ST			FORT LAU	FL	33312	5645	Y		3031 SW 47TH ST			FORT LAU	FL	33312	5645
12011	BROWAR	50-42-33-15-0120	294 NW 13TH CT			DANIA BE	FL	33004	2613	Y		294 NW 13TH CT			DANIA BE	FL	33004	2613

Prepped.xlsx



policyno	ZIPCODE	TOTEXPC	PREMEA	GROUP_I	SYSRAND	SYS_EEA	SYS_Polic	SYS_Polic	EffectiveD	ExpirationI	SYS_New	SYS_Ren	SYS_Polic	EEA_Polic	LPS_OutOfStateMail	LPS_Owne	LPS_Land
474231260080	33073	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	474231260	2013	N	Y	Residentia
474233080070	33442	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	474233080	2013	N	Y	Residentia
474235061050	33442	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	474235061	2013	N	Y	Residentia
484106150480	33076	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	484106150	2013	N	Y	Residentia
484107130690	33076	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	484107130	2013	Y	~	Residentia
484108023700	33076	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	484108023	2013	N	~	Residentia
484110160240	33067	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	484110160	2013	N	Y	Residentia
484111040720	33067	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	484111040	2013	N	Y	Residentia
484111061030	33067	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	484111061	2013	N	Y	Residentia
484115030139	33065	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	484115030	2013	N	Y	Residentia
484124020350	33063	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	484124020	2013	N	Y	Residentia
484125090400	33063	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	484125090	2013	N	Y	Residentia
484125210750	33063	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	484125210	2013	N	Y	Residentia
484127040480	33071	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	484127040	2013	N	Y	Residentia
484133030380	33071	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	484133030	2013	N	~	Residentia
484135041690	33063	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	484135041	2013	N	Y	Residentia
484135060930	33068	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	484135060	2013	N	~	Residentia
484214020340	33064	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	484214020	2013	N	Y	Residentia
484224150620	33064	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	484224150	2013	N	~	Residentia
484232100420	33066	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	484232100	2013	N	Y	Residentia
484306440151	33441	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	484306440	2013	N	Y	Residentia
484317031590	33064	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	484317031	2013	N	Y	Residentia
484318080230	33064	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	484318080	2013	N	~	Residentia
494024100970	33323	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	494024100	2013	N	Y	Residentia
494035031080	33323	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	494035031	2013	N	~	Residentia
494102090361	33068	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	494102090	2013	N	Y	Residentia
494104410140	33321	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	494104410	2013	Y	Y	Residentia
494105050720	33321	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	494105050	2013	N	Y	Residentia
494105060520	33321	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	494105060	2013	N	Y	Residentia
494106250690	33321	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	494106250	2013	N	~	Residentia
494108050920	33321	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	494108050	2013	N	~	Residentia
494110292290	33068	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	494110292	2013	Y	~	Residentia
494111190403	33068	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	494111190	2013	N	~	Residentia
494113031490	33319	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	494113031	2013	Y	Y	Residentia
494114110940	33319	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	494114110	2013	N	Y	Residentia
494115161360	33319	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	494115161	2013	Y	~	Residentia
494115192330	33319	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	494115192	2013	N	Y	Residentia
494117230390	33351	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	494117230	2013	N	Y	Residentia
494120300160	33351	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	494120300	2013	N	Y	Residentia
494126330130	33313	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	494126330	2013	N	Y	Residentia
494127270180	33322	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	494127270	2013	N	Y	Residentia
494128140840	33322	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	494128140	2013	N	Y	Residentia
494129210145	33322	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	494129210	2013	N	~	Residentia
494132030300	33322	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	494132030	2013	N	Y	Residentia
494212072250	33308	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	494212072	2013	N	Y	Residentia
494212082960	33308	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	494212082	2013	N	Y	Residentia
494215040330	33334	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	494215040	2013	N	Y	Residentia
494229511090	33311	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	494229511	2013	N	Y	Residentia
494232013110	33069	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	494232013	2013	N	~	Residentia
494306250200	33062	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	494306250	2013	N	Y	Residentia
504001370030	33325	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	504001370	2013	N	Y	Residentia
504105201570	33324	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	504105201	2013	N	Y	Residentia
504107170410	33324	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	504107170	2013	N	Y	Residentia
504108310080	33324	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	504108310	2013	N	Y	Residentia
504113032600	33317	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	504113032	2013	N	Y	Residentia
504114050230	33317	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	504114050	2013	N	Y	Residentia
504114140120	33317	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	504114140	2013	N	Y	Residentia
504114160510	33317	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	504114160	2013	N	Y	Residentia
504115040610	33317	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	504115040	2013	N	Y	Residentia
504121170640	33328	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	504121170	2013	N	Y	Residentia
504128350150	33328	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	504128350	2013	N	Y	Residentia
504129020200	33328	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	504129020	2013	N	Y	Residentia
504132101000	33328	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	504132101	2013	N	Y	Residentia
504132131280	33328	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	504132131	2013	N	Y	Residentia
504202150530	33301	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	504202150	2013	N	Y	Residentia
504205151150	33069	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	504205151	2013	N	~	Residentia
504208120210	33312	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	504208120	2013	N	~	Residentia
504211013130	33301	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	504211013	2013	N	Y	Residentia
504215260110	33315	1	100	2013	-1	100	20131211	20141211	20131211	20141211	Y	Y	504215260	2013	N	~	Residentia

Milliman

Review the prepped data

- Number of records received
- List of fields
- Sample of records
- Distributions of all variables – check for reasonable values
- Produce summary by year

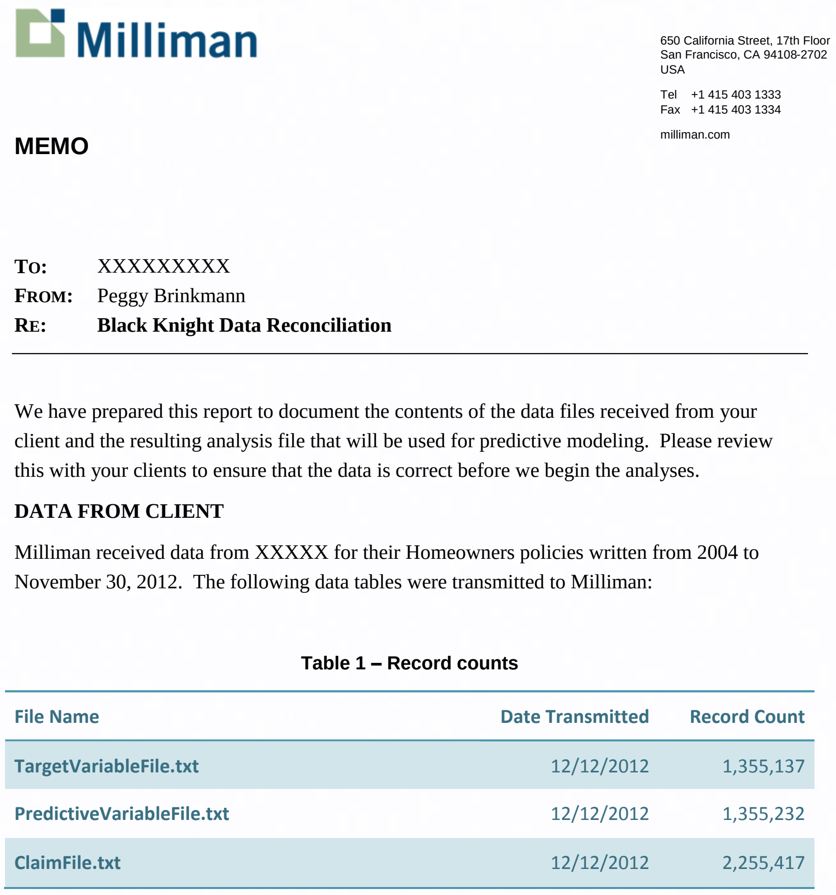
* CASE STUDY *

QC Prepped Data.txt

Documentation

- Raw data report
- Data layout
- Manual checks
- Prepped data report
- Reconciliation report (from raw to prepped)
  - Compare summaries, record counts
  - Document data adjustments (onlevel, trend, capping, etc.)
  - Document exclusions/filters (e.g. noncat loss only)

* CASE STUDY *



650 California Street, 17th Floor
San Francisco, CA 94108-2702
USA

Tel +1 415 403 1333
Fax +1 415 403 1334

milliman.com

MEMO

TO: XXXXXXXXX
FROM: Peggy Brinkmann
RE: **Black Knight Data Reconciliation**

We have prepared this report to document the contents of the data files received from your client and the resulting analysis file that will be used for predictive modeling. Please review this with your clients to ensure that the data is correct before we begin the analyses.

DATA FROM CLIENT

Milliman received data from XXXXX for their Homeowners policies written from 2004 to November 30, 2012. The following data tables were transmitted to Milliman:

Table 1 – Record counts

File Name	Date Transmitted	Record Count
TargetVariableFile.txt	12/12/2012	1,355,137
PredictiveVariableFile.txt	12/12/2012	1,355,232
ClaimFile.txt	12/12/2012	2,255,417

Questions?

THANK YOU

peggy.brinkmann@milliman.com